Jurnal Teknologi

Full paper

An Improved Method for Predicting and Ranking Suppliers Efficiency Using Data Envelopment Analysis

Mohammadreza Farahmand[a,b]*, Mohammad Ishak Desa[a,c], Mehrbakhsh Nilashi[a], Antoni Wibowo[a]

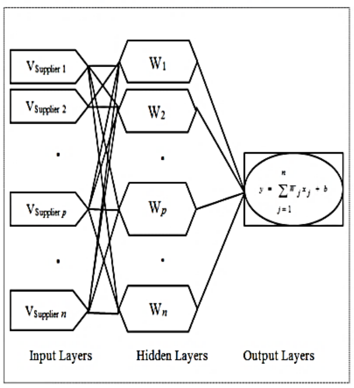
[a]Faculty of Computing, Universiti Teknologi Malaysia, 81310 UTM Johor Bahru, Johor, Malaysia
[b]Department of Science, Abarkouh Branch, Islamic Azad University, Abarkouh, Iran
[c]Advanced Informatics School, Universiti Teknologi Malaysia, Jalan Semarak, 54100 Kuala Lumpur, Malaysia

*Corresponding author: mmreza_farahmand@yahoo.com

Article history

Received :10 December 2014
Received in revised form :
1 February 2015
Accepted :12 February 2015

Graphical abstract



Abstract

Supplier selection problem (SSP) is a problem to select the best among suppliers based on input and output data of the suppliers. Since different uncontrollable and unpredictable parameters are affecting selection, choosing the best supplier is a complicated process. Data Envelopment Analysis (DEA) is a method for measuring efficiency and inefficiencies of Decision Making Units (DMUs). DEA has been employed by many researchers for supplier selection and widely used in SSP with inputs for supplier evaluation. However, the DEA still has some disadvantages when it is solely used for SSP. Hence, in this paper, a combination of DEA and Neural Network (NN), DEA-NN, is proposed for SSP. We also develop a model for SSP based on Support Vector Regression (SVR) to improve the stability of DEA-NN. The proposed method was evaluated using small and large data sets. The experimental results showed that, the proposed method solve the problems connected to the previous methods. The results also showed that stability of proposed method is significantly better than DEA-NN method. In addition, CCR-SVR model overcome shortcomings such as instability and improves computational time and accuracy for predicting efficiency of new small and large DMUs.

Keywords: Supplier selection problem; data envelopment analysis; neural networks; support vector machines; support vector regression; decision making units

1.0 INTRODUCTION

Supplier selection problem (SSP) is a problem to select the best among suppliers based on input and output data of the suppliers. Selection of the most appropriate suppliers is based on their ability to meet some criteria such as providing the right products for buyers, quality of services at the right time, price and quantities.[1] Since different uncontrollable and unpredictable parameters are affecting selection, choosing the best supplier is a complicated process. Various decision making approaches have been proposed to tackle the problem in contemporary supply chain management. The performance of potential suppliers is evaluated against multiple criteria rather than considering a single factor-cost.[2] The multi-criteria decision making approaches are better than the traditional cost-based approaches, which aids the researchers and decision makers in applying the approaches effectively.[3] Several methods including Analytic Hierarchy Process (AHP), Analytic Network Process (ANP), Case-Based Reasoning (CBR), Fuzzy Set Theory (FST), Genetic Algorithm (GA), Mathematical Programming (MP) Models, Simple Multi-Attribute Rating Technique (SMART), Data Envelopment Analysis (DEA) and combination of the methods have been proposed for SSP.[4-8, 9] The DEA method, which is a linear programming technique for computing the efficiency DMUs, was first developed.[10] DEA application in supplier evaluation has been presented.[11-16] It has been widely used for efficiency estimation in both private and public sections of organizations such as banks, hospitals, airlines, universities, etc. and attract a great deal of researchers attention because of its ability in performance assessment. Recently, individual of DEA approaches have been presented for evaluation and selection process.[17] Also, supplier development strategies with DEA have been proposed.[18] While, the DEA has been employed by several researchers but it still has disadvantages which are expressed as the following: a) inefficient units can be ranked according to their inefficiency, but efficient units (units with efficiency equal to 1) cannot be ranked. Also, with increasing number of input and output, the number of efficient units will increase. Obviously, without rank, it is not possible to choose the best among the efficient units, b) DEA for a large dataset with many inputs or outputs would require an advanced computer with high resources in terms of memory and processing speed, c) with addition of new units to old units, DEA cannot obtain efficiency of new units without recalculating the efficiency of all the units. In facts, DEA cannot be useful for predicting the efficiency new units.

Therefore, an integration of NN and DEA for supplier evaluation has been proposed based on incomplete information of criteria.[8] In their work, they highlight that DEA often fails to work effectively because DEA is associated with homogeneity and accuracy assumption. An integrated model of NNs and DEA has also been proposed for supplier evaluation system.[19] They commented DEA still requires more compute resources in comparison with NNs which are good tool for simulation and learning. It is emphasized that developing NN can be useful to reduce the time of DEA process. In addition it paves the way for prediction or checking of suppliers' efficiency. An integrated model of DEA, Decision Tree (DT) and NNs has been presented for evaluation of supplier's performance based on several models including efficient and inefficient clusters using the result of DEA and application of DT and NN.[20] An integration of NNs and DEA has been proposed for measuring the efficiency of large collections with model of output oriented in DEA.[21] An application of DEA with a factor of undesirable has been applied for evaluation of branch efficiency in the Taiwanese bank.[22]

A comparative study of supplier selection based on SVM and Radial Basic Function Neural Network (RBFNN) also showed that SVM is more superior, giving more accurate results, than RBFNN algorithm.[23] In addition, a new combined method was proposed using DEA and Support Vector Regression (SVR), DEA-SVR, for efficiency evaluation of large DMUs to solve some drawbacks which include uncontrolled convergence and non-generalization.[24]

As can be found from the above literature, DEA has been widely used for applications with inputs and outputs; however, the approaches used in the prior researches have some drawbacks which include model infeasibility due to too small or too big values of input, suffer from obtaining complete ranking and non-generalization. Thus, in this paper, the combination of DEA and NNs is adopted and extended to evaluate different suppliers from two essential aspects, namely efficiency prediction and ranking. It is shown that the combined model can overcome the previously mentioned drawbacks of DEA. In addition, the accuracy and stability in measuring the efficiency of suppliers for ranking and classification still have been two main challenges in the field. In addition, the approaches used in the prior researches have some drawbacks which include time consuming and insufficient accuracy.

We also improved stability of prediction algorithm for the prediction method which is discussed in this study. In this paper, accordingly, predicting efficiency of the new supplier and stability analysis of the proposed methods are examined for DEA-NN. In addition, we show the effect of parameters of proposed method on stability analysis, two examples applying on small and large data sets are considered in the following for proposed algorithm.

The remainder of this paper is organized as follows. Section 2 provides research methodology and hybrid proposed model. Section 3 presents the experimental results and performance comparisons and finally, conclusions are presented in Section 4.

2.0 METHODOLOGY

One of the most popular NN algorithms is Back Propagation (BP) algorithm. Rojas, (1996)[25] claimed that BP algorithm could be divided to four main phases such as feed-forward computation, BP to the output layer, back propagation to the hidden layer and Weight updates. After choosing the weights of the network randomly, the BP algorithm is used to compute the necessary corrections. The learning processing of BP Neural Network is divided to three main processes such as training, validation and testing. Training works for to estimate the weights and bias of network design selection and validation is estimated for performance error of network design which is needed to stop training. The third process is obtained for the final weights and bias. Time to train NN is probably identified as biggest disadvantage. The Back-Propagation algorithm can be applied on multilayer feed-forward NNs and perform the learning.[26]

The Back-Propagation algorithm can be applied on multilayer feed-forward NNs and perform the learning. The data used for training the NN includes input, output and efficiency of DEA.[19,21]

2.1 The Proposed Method

We consider input and output of DEA for suppliers as the neural network input and the efficiency as the output of neural networks. The efficiencies of suppliers are computed using CCR and BCC models and then use them as input to DEA-NN method. The Back-Propagation NN in DEA is trained by iteratively processing a training sample, comparing the network's prediction of efficiency scores for each sample of DMUs with actual known efficiency scores. In this section two traditionally used methods, namely DEA and NN, and the combination of the two (DEA-NN) for measuring efficiency DMUs classification and ranking are presented. In the proposed model combine two common methods of DEA, namely CCR and BCC, with NNs. The models call CCR-NN and BCC-NN respectively for the CCR and BCC models combined with NNs. The proposed models are used for ranking and efficiency prediction of DMUs. To evaluate the accuracy of the proposed models, the ranking data obtained from DEA and NNs are compared with CCR-NN and BCC-NN models. The data used for training the neural network includes input, output and efficiency of DEA. The Back-Propagation NN in DEA is trained by iteratively processing a training sample, comparing the network's prediction of efficiency scores for each sample of DMUs with actual known efficiency scores. Figure 1 shows the Back-Propagation application in DMUs method and this combined method is called DEA-NN method.

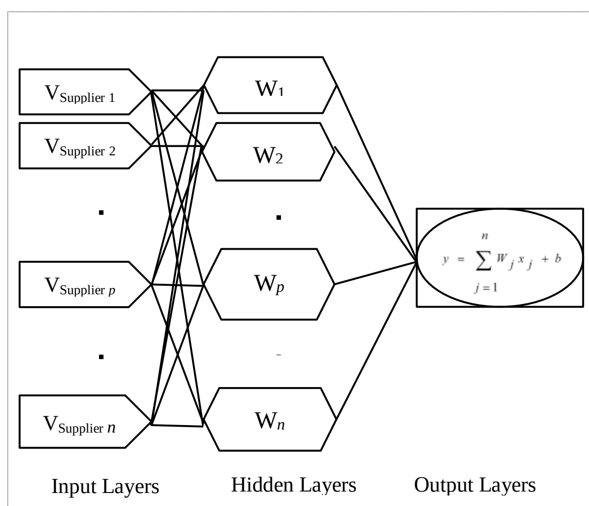


Figure 1 A DEA-NN algorithm for predicting efficiency of suppliers set, where $V_{Supplier\ p}$ and W_p are input vector in input layers and weight in hidden layers for pth supplier

In this model, assume that are n DMUs with $X=[X_1,..., X_n]$ and $Y=[Y_1,...,Y_n]$. Vectors X_j and Y_j denote the observed data where, $X_j=(x_{1j},x_{2j},\ldots,x_{mj})$ is a column vector of observed inputs and $Y_j=(y_{1j},y_{2j},\ldots,y_{sj})$ is a column vector of observed outputs for each DMU.

Figure 2 shows the DEA-NN algorithm application in DMUs method and this combined method is called CCR-NN method. It shows the general flow of steps in using CCR and CCR-NN for DMUs' efficiency prediction and ranking. In addition, process of DEA-NNs method improved in Algorithm 1 as following:

Algorithm 1: Process of DEA-NNs method

Step 1 (Data collection): In this step input data are based on inputs of the large DMUs. Suppose $X=[x_{ij}]_{m\times n}$ and $Y=[y_{rj}]_{s\times n}$ are input and output data respectively. Notice that, commonly the number of evaluation units is more or equal than three times the total number of inputs and outputs data i.e. $n \geq 3(m+s)$.

Step 2 (Calculate DMU efficiency using CCR/BCC model): The efficiency of DMUs are determined using CCR and BCC models with LP algorithm for each DMU, where Θ_j is the efficiency of DMU$_j$ (j=1, 2, …, n).

Step 3 (Training of data with CCR-NN/BCC-NN): The training algorithm (CCR-NN/BCC-NN) is initialize training *epoch*=1 (epoch is a step in the training process) and initialize weights and biases with random values. Import to input data of network and calculate output data. (In this study $[X|Y]_{M\times n}$ ($M=m+s$) and $[\Theta]$ are network input data and network target data respectively). If $MSE \leq MSE_{min}$ then stop training network (where *MSE* is the network Mean Squares Errors which is used for training of the network). If $epoch \geq epoch_{max}$ then stop training otherwise update weights and biases and $epoch=epoch+1$ and go to Step3.

Step 4 (Are results satisfactory?): Handle and select the best of data based on the performance during training, testing and validation, and if the results are satisfactory then go to the next step otherwise go to Step 3.

Step 5 (Obtain efficiency of DMU based on CCR-NN/BCC-NN): Simulate input and target data based on the training data from Step 4 and obtain efficiency of DMUs.

Step 5a.1 (Get new DMUs): Collect new data based on Step 1 (in this study $X_{\text{new}}=[x_{ij}]_{m\times n^*}$, $Y_{\text{new}}=[y_{rj}]_{s\times m}$ and $[X_{\text{new}} \mid Y_{\text{new}}]_{M\times n^*}$ is new input data of network, where n^* is number of new units).

Step 5a.2 (Prediction of performance/efficiency): Simulate $[X_{\text{new}}|Y_{\text{new}}]_{M\times n^*}$ based on training of $[X|Y]_{M\times n}$ and $[\Theta]$ from results of Step 4 for CCR-NN/BCC-NN and obtain $[\Theta_{\text{new}}]$.

Step 5b.1 (Ranking of DMUs): Rank $[\Theta]$ and $[\Theta_{\text{new}}]$ based on result of Step 5.

Step 6 (Comparison between results of models): Comparison between results of Step 5a.2 and 5b.1

Step 7: Stop.

In the obtained results of CCR and BCC models, one can find some suppliers with the same efficiency. Therefore suppliers with equal efficiency cannot be ranked by CCR and BCC models. To overcome this problem, algorithms were developed which is a combination of CCR with NN (CCR-NN) and BCC with NN (BCC-NN).

In addition, the so called CCR-NN based efficiency prediction (CCR-NNP), and BCC-NN based efficiency prediction (BCC-NNP) models were also developed for prediction process of new suppliers. The predicted results obtained from these two models are then compared against each other. We also are considered a classification with 2 class class-classification that supplier to be classified as "Almost Efficient (AE)", if the efficiency is less than 0.001 and otherwise supplier is "Inefficient (IE)".

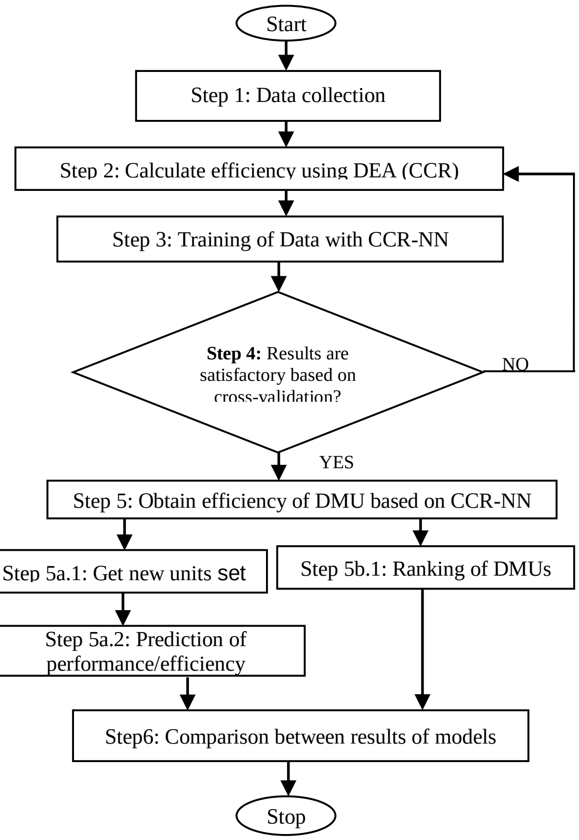


Figure 2 CCR-NN algorithm for efficiency prediction by DEA (CCR)-NN model

2.3 Improving the DEA-NN using stability analysis with Support Vector Regression (SVR)

Accuracy and computational time as two main issues were considered for testing the new proposed methods. The proposed combination of DEA and SVR (DEA-SVR) method was proposed in[24] for predicting efficiency of large DMUs. Thus, stability of prediction algorithm is one of another important issue for prediction methods which is discussed in this section. Accordingly, predicting efficiency of the new DMUs and stability analysis of the for DEA-NN method are examined for DEA-SVR.[24] We are considered Training (Tr) and testing (Te) sets for testing stability analysis of predicting efficiency of large DMUs in Table 1. In addition, this division is randomly selected by user.

Table 1 Conditions for selection training and testing sets in large DMUs

Conditions	D=Data sets	T_r = 80% D	T_e = 20% D

To show the effect of parameters of proposed methods on stability analysis, two examples applying on small and large data sets are considered in the following for proposed algorithm.

Algorithm 2. Procedure of stability analysis for DEA-NN\SVR method

Step 1: For *i=1 to n* Do (*n* is number of trial)
Step 2. $j \leftarrow i$.
Step 3: Obtain data set collection for input NN\SVR (*D*).
Step 4: Divided *D* set to two blocks such as 80 % and 20% from *D* set for training (T_r) and testing (T_e) sets respectively (Table 1).
Step 5: Obtain new units set (*D**).
Step 6: Predict efficiency of *D** based on Step 2.
Step 7: Calculate Standard Deviation (SD) for each trail.
Step 8: If $j<n$ go to step 1 else go to Step 8.
Step 9: Stop.

3.0 EXPERIMENTAL RESULTS AND PERFORMANCE COMPARISONS

This section explains the experiments carried out to test the performance of the proposed DEA-SVR algorithm for large DMU's data sets. The performance is in terms of accuracy which is measured by several evaluation methods. Performance comparisons between normalization functions, DEA-SVR models, predicting efficiency of new DMUs and stability analysis of DEA-SVR with DEA-NN are also presented.

3.1 Evaluation Methods

In this section, the experiments perform on large data sets to test the performance of the proposed method. The performance is defined in terms of accuracy which is measured by reduction of efficient units to obtain the DMU's ranking. In our experiments, performance comparisons between proposed method and CCR model and integrated DEA with prediction methods are presented. For evaluating the proposed method, two measures of accuracy are used to determine the algorithm capability for DMUs ranking.

a) *K*-Fold cross-validation: *K*-Fold cross-validation estimates of performance cross-validation is a computer intensive technique, using all available examples as training and test examples. It makes pattern for the use of training and test sets by repeatedly training the algorithm *K* times with a fraction $1/K$ of training examples left out for testing purposes. This kind of hold-out estimate of performance lacks computational efficiency due to the repeated training, but the latter are meant to lower the variance of the estimate. One of the cross validation metric is used based on as following[27]:

Mean Squared Error (MSE):

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(f(x_i) - y_i)^2 \tag{1}$$

where $f(x_i)$ must be approximated by y_i.
b) Correlation test: correlation coefficient is measuring of relation between two variables such as *x* and *y*, and value of this measuring is between –1 and 1. If the two variables are in perfect linear relationship, the correlation coefficient will be either 1 or –1. The sign depends on whether the variables are positively or negatively related. If correlation coefficient is 0 then there is no linear relationship between the variables. Suppose (x_1,y_1), (x_2,y_2), . . . ,(x_n,y_n) are paired measurements, the Pearson productmoment correlation coefficient is given in [28]:

$$r_{xy} = CC = \frac{(n\sum_{i=1}^{n} x_i y_i - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i)}{\sqrt{(n\sum_{i=1}^{n} x_i^2 - (\sum_{i=1}^{n} y_i)^2)(n\sum_{i=1}^{n} y_i^2 - (\sum_{i=1}^{n} y_i)^2)}} \tag{2}$$

where i=1,2,…,n and *CC* is correlation coefficient for x_i and y_i. Also Squared Correlation Coefficient (*SCC*) for x_i and y_i as follow [28]:

$$SCC = \frac{(n\sum_{i=1}^{n} x_i y_i - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i)^2}{(n\sum_{i=1}^{n} x_i^2 - (\sum_{i=1}^{n} y_i)^2)(n\sum_{i=1}^{n} y_i^2 - (\sum_{i=1}^{n} y_i)^2)} \tag{3}$$

where *SCC* is between 0 and 1.
c) Accuracy Evaluation Formula:
One calculates the classification accuracy using the following equation. [29]

$$\text{Accuracy }(\%) = \frac{\text{Correctly predicted data}}{\text{Total testing data}} \times 100 \tag{4}$$

3.2 Data Set Description

The experimental data set for this experiment was taken from[18] which consist of 25 suppliers as in [18] was used. Table 2 shows the sample includes 2 Inputs (I_1: Quality management Practices & I_2: Employee training) and 3 Outputs (O_1: Quality of the product, O_2: Price of the product and O_3: Delivery of the product), where I_i and O_r are inputs and outputs data, respectively (i=1,2 and r=1,2,3).

Table 2 Sample of large data set 1 with 25 units, 2 inputs and 3 outputs

Suppliers	I_1	I_2	O_1	O_2	O_3
Supplier 1	73	99	60	50	35
Supplier 2	45	67	87	45	50
Supplier 3	78	87	43	35	60
….	….	….	….	….	….
Supplier 25	78	65	65	56	70

The experimental data set for this experiment was taken from[21] which consist of 5 data sets. Each data set contains 5000 units and each unit has 6 attributes of 3 inputs and 3 outputs. A sample of original data set 1 is shown in Table 3. The program code for DEA-SVR and DEA-NN (i.e. CCR-NN) are improved by LIBSVM [28] implemented in MATLAB software.

Table 3 Sample of large data set 1 with 5000 units, 3 inputs and 3 outputs

Units	I_1	I_2	I_3	O_1	O_2	O_3
U1	971	99	471	1777	575	7284
U2	190	12	4763	3312	184	2884
U3	732	83	9019	96	1935	8575
..….	……	……	……	……	……	……
U5000	555	34	3686	4965	857	1769

In this paper, six different supplier selection models, CCR, BBC, CCR-NN, BBC-NN, CCR-NNP and BCC-NNP, were compared and evaluated based on the obtained efficiency measure of suppliers.

3.3 Experimental Results Using DEA-NN for Predicting Efficiency of New Suppliers

The predicting efficiency of efficient suppliers (CCR-efficiency and BCC-efficiency) obtained using CCR-NN and BCC-NN models are shown in Table 4 results of the experiments should be described and discussed in this section.

The ranking obtained by CCR and CCR-NN models are compared for efficient suppliers (efficiency equal to one) in Table 4. Based on CCR model alone, 4 suppliers (Supplier 2, 4, 11, and

16) have the same ranking (any rank between 1 to 4). With CCR-NN, on the other hand, it is possible to rank the best 4 suppliers and in this case Supplier 4 is the best supplier.

Table 4 Comparison of ranking between efficient suppliers with CCR and CCR-NN model

Suppliers	Efficiency (CCR)	Efficiency (CCR-NN)	Ranking (CCR-NN)
Supplier 2	1	0.995619129	3
Supplier 4	1	0.99998511	1
Supplier 11	1	0.999985061	2
Supplier 16	1	0.980550588	4

Similarly when the BCC model is used, it shows that the BBC model alone is unable to give definite ranking to individual supplier. In fact, it is more indecisive here since, as shown in Table 5, the number of suppliers with seemingly equal ranking is larger (that is 9 suppliers compared to 4 suppliers with CCR model). With BCC-NN, 9 suppliers can be distinctively ranked.

Table 5 Comparison of ranking between efficient suppliers with BCC and BCC-NN model

Suppliers	Efficiency (BCC)	Efficiency (BCC-NN)	Ranking (BCC-NN)
Supplier 2	1	0.998778551	6
Supplier 4	1	0.999771368	3
Supplier 7	1	0.766661621	9
Supplier 10	1	0.99977623	1
Supplier 11	1	0.999756075	2
Supplier 16	1	0.999229412	5
Supplier 17	1	0.975189832	8
Supplier 21	1	0.999493	4
Supplier 24	1	0.998732389	7

It can be seen that CCR-NN and BCC-NN models outperform the traditional models in the accuracy of the efficiency estimation. The predicted data for five new suppliers obtained from CCR-NNP and BCC-NNP are shown in Table 6. It can be seen that five new suppliers were ranked using the models.

Table 6 Efficiency and ranking with models of CCR-NNP and BCC-NNP for new suppliers

New Suppliers	Efficiency (CCR-NNP)	Rank (CCR-NNP)	Efficiency (BCC-NNP)	Rank (BCC-NNP)
Supplier 26	0.99999347	1	0.999788667	3
Supplier 27	0.999860827	2	0.999829813	1
Supplier 28	0.596543046	5	0.690622256	5
Supplier 29	0.998448658	3	0.999827928	2
Supplier 30	0.995619129	4	0.998778551	4

Table 7 shows the comparison of classification for 5 new suppliers using CCR-NNP and BCC-NNP with two class AE and IE that are almost efficient and inefficient suppliers.

Table 7 Classification with models of CCR-NNP and BCC-NNP for new Suppliers (AE and IE are almost efficient and inefficient class respectively)

New Suppliers	Classification (CCR-NNP)	Classification (BCC-NNP)
Supplier 26	AE	AE
Supplier 27	AE	AE
Supplier 28	IE	IE
Supplier 29	IE	AE
Supplier 30	IE	IE

To design our network, we adopted training iteration for the optimum number of hidden nodes in hidden layers. In our simulation, 4 and 8 hidden layers are considered. To design the optimum NNs, we used the iteration method of training to obtain the optimum numbers of layers and nodes.

3.4 Experimental Results Using Stability Analysis Between DEA-SVR and DEA-NN Methods for Predicting Efficiency Of New DMUs

In the following sections, experimental results using stability analysis are obtained using DEA-SVR and DEA-NN methods for predicting efficiency of new DMUs. The experiments for standard deviation, accuracy and run time of Algorithm 2 are provided. Thus, two types of data sets have been used for proposed algorithm evaluation. The first one is small data set taken from[18] (see Table 2) and the second data set is large DMUs set which has been taken from[21] (see Table 3).

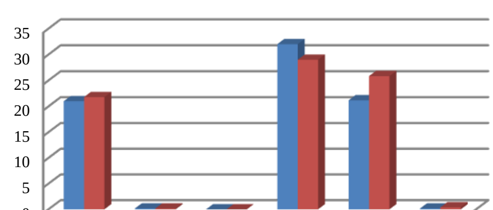
a) Example 1 (small data set): This section presents the results of comparison using DEA-SVR and DEA-NN methods based on Mean Squared Error (MSE), standard deviation (SD) and computational time for predicting efficiency of new small suppliers with testing stability analysis of methods DEA-SVR and DEA-NN in Algorithm 2.

Table 8 shows the results of comparison between CCR-SVR and CCR-NN with conditions 1-3. Condition 1 for first model (CCR-SVR) is included C=7, γ= 2 and v=0.1 and 100 trials and for second model (CCR-NN) are 10 Hidden Layers (HL) with 100 trials. Condition 2 for first model are included C=6, γ= 2 and v=0.1 and 100 trials and for second model are 50 HL with 100 trials. Condition 3 for first model are included C=8, γ= 2 and v=0.1 and 100 trials and for second model are 100 HL with 100 trials. From this table, it can be seen that for CCR-SVR and CCR-NN models the percentage of the MSE score, SD and time total are obtained for conditions 1-3. Average of standard deviation score is obtained 0 and 0.2836 for CCR-SVR and CCR-NN respectively. Thus, CCR-NN is instability and with using CCR-SVR for predicting efficiency of new suppliers, computational time and accuracy have obtained suitable result. In Table 8, averages of MSE using first and second model are obtained 0.002146 and 0.00846, respectively. In addition, from this table, average of total time using first and second models are obtained 0.165 and 22.678 second, respectively.

Table 8 A comparison between CCR-SVR and CCR-NN models for stability analysis based on accuracy, standard deviation (SD) and computational time (total time) using predicting efficiency of new suppliers

Models	Conditions		
	1	2	3
CCR-SVR			
(C,γ)	(7,2)	(6,2)	(8,2)
Trial	100	100	100
MSE	0.00216	0.0021	0.00218
SD	0	0	0
Total Time (Second)	0.1620	0.1740	0.1590
CCR-NN			
Hidden Layers(HL)	10	50	100
Trial	100	100	100
MSE	0.0362	0.0032	0.0029
SD	0.0332	0.1729	0.3943
Total Time (Second)	21.0430	21.1780	25.8410

Figure 3 shows the results of comparison between CCR-SVR and CCR-NN for accuracy, computational time and stability for conditions 2 and 3. From this figure, it can be seen that for CCR-SVR the percentage of MSE score is obtained 21% and 21.8% for conditions 2 and 3 respectively. Also, for CCR-NN, the percentage of the MSE score is obtained 32% and 29% for conditions 2 and 3 respectively. Averages of SD score are obtained 0 and 0.2836 for CCR-SVR and CCR-NN respectively. Hence, it can be concluded that CCR-SVR is more stable than CCR-NN model. Also, computational time for CCR-SVR is very less than CCR-NN.



	MSE %(CCR-SVR)	Time (second), CCR-SVR	SD(CCR-SVR)	MSE %(CCR-NN)	Time(second), CCR-NN	SD(CCR-NN)
Condition 2	21	0.174	0	32	21.178	0.1729
Condition 3	21.8	0.159	0	29	25.841	0.3943

Figure 3 Comparison between CCR-SVR and CCR-NN models with conditions 2 and 3 for new small suppliers with 100 trial

It is clear from this results that the CCR-SVR model shows its superiority over CCR-NN model for predicting efficiency of new small suppliers.

b) Example 2 (large data set): The data set used in [21] as a large data set has three inputs and three outputs in the five groups indicating data sets 1-5 (Section 3.2).

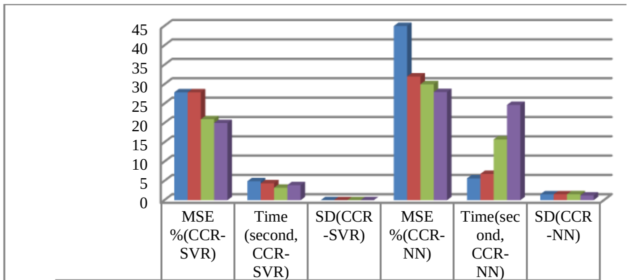
This section presents the results of comparison between DEA-SVR and DEA-NN methods using MSE, SD, Squared Correlation Coefficient (SCC) and Time Average (TA (second)) for predicting efficiency of new large DMUs in testing stability analysis of CCR-SVR and CCR-NN models.

Table 9 shows a comparison between CCR-SVR and CCR-NN models with conditions 1-4. Condition 1 for first model (CCR-SVR) are included C=10, γ= 7 and ν=0.08 and 10 trials and for second model (CCR-NN) are 10 Hidden Layers (HL) with 10 trials. Condition 2 for first model are included C=9, γ= 7 and ν=0.08 and 10 trials and for second model are HL with 10 trials. Condition 3 for first model are included C=6, γ= 8 and ν=0.08 and 10 trials and for second model are 50 HL with 10 trials. Condition 4 for first model are included C=10, γ= 7.1 and ν=0.08 and 10 trials and for second model are 100 HL with 10 trials. Average of SD score is obtained 0 and 1.4525 for CCR-SVR and CCR-NN respectively. Thus, CCR-NN model is instability with comparing the CCR-SVR model for predicting efficiency of new large DMUs. In Table 9, average of MSE using first and second models are obtained 0.00242075 and 0.003375, respectively. In addition, from this table, average of time average using first and second models are obtained 4.1075 and 13.22 second, respectively. Furthermore, averages of SCC using first and second models are obtained 0.935 and 0.86, respectively. From this table, it can be seen, computational time and accuracy have obtained suitable result for CCR-SVR model.

Table 9 A comparison between DEA-SVR and DEA-NN based on MSE, Standard Deviation (SD) and Time Average (TA) using prediction of new large DMUs for stability with CCR-SVR model for DATA 1-5

Models	Conditions			
	1	2	3	4
CCR-SVR				
(C,γ)	(10,7)	(9,7)	(6,8)	(10,7.1)
Trial	10	10	10	10
MSE	0.00279	0.00279	0.0021	0.0020
SD	0	0	0	0
SCC	0.93004	0.92981	0.9485	0.9501
TA (Second)	4.9290	4.3830	3.2290	3.8980
CCR-NN				
Hidden Layers(HL)	10	20	50	100
Trial	10	10	10	10
MSE	0.0045	0.0032	0.003	0.0028
SD	1.52	1.53	1.54	1.22
SCC	0772	0.871	0.8902	0.916
TA (Second)	5.6606	6.802	15.748	24.6836

Figure 4 shows the results of comparison between CCR-SVR and CCR-NN using MSE, TA, SCC and SD for conditions 1-4. From this figure, it can be seen that for CCR-SVR the percentage of the MSE score is obtained between 20 to 27.92 % for conditions 1-4. Also, for CCR-NN model, the percentage of the MSE score is obtained between 28 to 45% for conditions 1-4. These results are generated using 5-Fold cross-validation for both of models. Average of SD score is obtained 0 and 1.4525 for CCR-SVR and CCR-NN respectively for testing stability. Hence, it can be concluded that CCR-SVR is more than stable CCR-NN. Also, computational time for CCR-SVR is very less than CCR-NN that TA scores are obtained 4.10975 and 13.22357 second for models of CCR-SVR and CCR-NN respectively. As a result, it is clear from this table that the CCR-SVR model shows its superiority over CCR-NN model for predicting efficiency of new large data sets.



	MSE %(CCR-SVR)	Time (second, CCR-SVR)	SD(CCR-SVR)	MSE %(CCR-NN)	Time(second, CCR-NN)	SD(CCR-NN)
Conditions 1	27.92	4.929	0	45	5.6606	1.52
Conditions 2	27.91	4.383	0	32	6.802	1.53
Conditions 3	21	3.229	0	30	15.7481	1.54
Conditions 4	20	3.898	0	28	24.6836	1.22

Figure 4 Comparison between CCR-SVR and CCR-NN based on conditions 1-4 for prediction of new DMUs with 10 trial

■4.0 CONCLUSION

In this paper, an improved method has been proposed for predicting and ranking in Supplier Selection Problem (SSP). We incorporated DEA and SVR in the proposed method to improve DEA-NN method. The proposed method has been evaluated using small and large data sets. The proposed method overcomes the problems connected to the DEA-NN method. The results also showed that stability of proposed method is significantly better than DEA-NN method. In addition, CCR-SVR model overcome shortcomings such as instability, computational time and insufficient accuracy for predicting efficiency of new small and large data sets.

References

[1] Zhang, D., Zhang, J., Lai, K.-K., and Lu, Y. (2009). An Novel Approach to Supplier Selection Based on Vague Sets Group Decision. *Expert Systems with Applications*. 36(5): 9557–9563.

[2] Ha, S. H., and Krishnan, R. 2008. A Hybrid Approach to Supplier Selection for the Maintenance Of A Competitive Supply Chain. *Expert Systems with Applications*. 34(2): 1303–1311.

[3] Kontis, A. p., and vrysagotis, V. 2011. Supplier Selection Problem: A Literature Review of Multi-criteria Approachs based on DEA. *Advances in Management & Applied Economics*. 1(2): 207–219.

[4] Aissaoui, N., Haouari, M., and Hassini, E. 2007. Supplier Selection and Order Lot Sizing Modeling: A Review. *Computers & Operations Research*. 34(12): 3516–3540.

[5] Akdeniz, H. A., and Turgutlu, T. 2007. Supplier Selection and Retail: Analysis with Two Multi Criteria Evaluation Methodologies. *Review of Social, Economic & Business Studies*. 9–10: 11–28.

[6] Amin, G. R., and Emrouznejad, A. 2011. Optimizing Search Engines Results Using Linear Programming. *Expert Systems with Applications. 38*(9): 11534–11537.

[7] Cebi, F., and Bayraktar, D. 2003. An Integrated Approach for Supplier Selection. *Logistics Information Management*. 10(6): 395–400.

[8] Çelebi, D., and Bayraktar, D. 2008. An Integrated Neural Network and Data Envelopment Analysis for Supplier Evaluation Under Incomplete Information. *Expert Systems with Applications*. 35(4): 1698–1710.

[9] Cooper, W. W., Seiford, L. M., & Tone, K. 2007. *Data envelopment Analysis: A Comprehensive Text with Models, Applications, References and DEA-Solver Software*. Second Editions. Springer, ISBN, 387452818, 490.

[10] Charnes, A., Cooper, W. W., and Rhodes, E. 1978. Measuring the Efficiency of Decision Making Units. *European Journal of Operational Research*. 2(6): 429–444.

[11] Weber, C. 1996. A Data Envelopment Analysis Approach to Measuring Vendor Performance. *Supply Chain Management: An International Journal*. 1: 28–39.

[12] Weber, C., Current, J., and Benton, W. 1991. Vendor Selection Criteria and Methods. *European Journal of Operations Research*. 50: 2–18.

[13] Weber, C., Current, J., and Desai, A. 1998. Non-cooperative Negotiation Strategies for Vendor Selection. *European Journal of Operations Research*. 108: 208–223.

[14] Weber, C., and Desai, A. 1996. Determination of Paths to Vendor Market Efficiency Using Parallel Co-Ordinates Representation: A Negotiation Tool for Buyers. *European Journal of Operations Research*. 90: 142–155.

[15] Weber, C., and Ellram, L. 1992. Supplier Selection Using Multi-Objective Programming: A Decision Support System Approach. *International Journal of Physical Distribution and Logistics Management*. 23(2): 3–14.

[16] Weber, C. A., Current, J. R., and Desai, A. 2000. An Optimization Approach to Determining the Number of Vendors to Employ. *Supply Chain Management: An International Journal*. 5(2): 90–98.

[17] Rumelhart, D. E., Hinton, G. E., and Williams, R. J. 1986. Learning Representations by Back-Propagating Errors. *Nature*. 323: 533–536.

[18] Mishra, R. K., & Patel, G. 2010. Supplier Development Strategies: A Data Envelopment Analysis Approach. *Business Intelligence Journal*. 3(1): 99–110.

[19] Dilek, Ö., and Gül, T. T. 2009. DEA ANN Approach in Supplier Evaluation System. *World Academy of Science, Engineering and Technology*. 54: 343–348.

[20] Desheng, W. 2009. Supplier Selection: A Hybrid Model Using DEA, Decision Tree and Neural Network. *Expert Systems with Applications*. 36: 9105–9112.

[21] Emrouznejad, A., and Shale, E. 2009. A Combined Neural Network and DEA for Measuring Efficiency of Large Scale Datasets. *Computers & Industrial Engineering*. 56(1): 249–254.

[22] Chang, K. C., Lin, C. L., Cao, Y., and Lu, C. F. 2011. Evaluating Branch Efficiency of a Taiwanese Bank Using Data Envelopment Analysis with an Undesirable Factor. *African Journal of Business Management*. 5(8): 3220–3228.

[23] Kong, Z. J., & Xue, J. B. 2013. January. A Comparative Study of Supplier Selection Based on Support Vector Machine and RBF Neural Networks. In *International Asia Conference on Industrial Engineering and Management Innovation (IEMI2012) Proceedings*. Springer Berlin Heidelberg. 917–926.

[24] Farahmand, M., Desa, M. I., & Nilashi, M. A. 2014. Combined Data Envelopment Analysis and Support Vector Regression for Efficiency Evaluation of Large Decision Making Units. *International Journal of Engineering and Technology (IJET)*. 2310–2321.

[25] Rojas, R. 1996. *Neural Networks: A Systematic Introduction*. Springer.

[26] Rumelhart, D. E., Hinton, G. E., & Williams, R. J. 1988. Learning Representations by Back-propagating Errors. *Cognitive Modeling*.

[27] Bengio, Y., & Grandvalet, Y. 2004. No Unbiased Estimator of the Variance of K-fold Cross-validation. *The Journal of Machine Learning Research*. 5: 1089–1105.

[28] Salkind, N. J. 2006. *Encyclopedia of Measurement and Statistics*. Sage Publications.

[29] Chang, C. C., & Lin, C. J. 2011. LIBSVM: A Library for Support Vector Machines. *ACM Transactions on Intelligent Systems and Technology (TIST)*. 2(3): 27.